Automating Life Detection Using Lipid Detection in GCMS Data

P. Michael Furlong, Mary Beth Wilhelm, Denise Buckner, Morgan Anderson, Matt Chin, Tori Chinn, Tom McClure, Kanch Sridhar, Travis Boone, Jennifer Eigenbrode, Trey Smith, Tony Ricco, George Cooper, Linda Jahnke

9 February 2021

NASA SILICON VALLEY
AMES RESEARCH CENTER
NASA Goddard SPACE FLIGHT CENTER
Shell

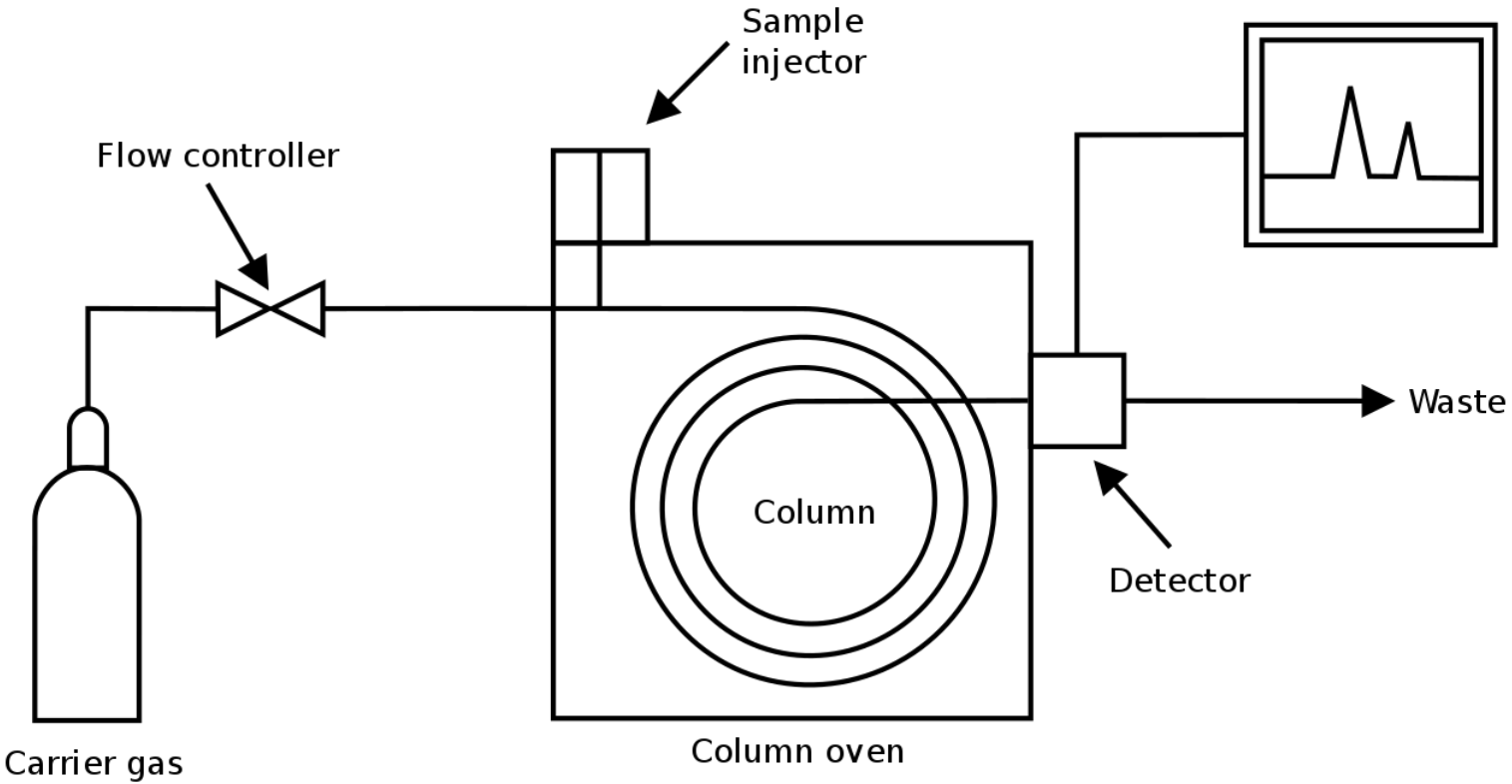
rough draft of project patch!

ExCALiBR

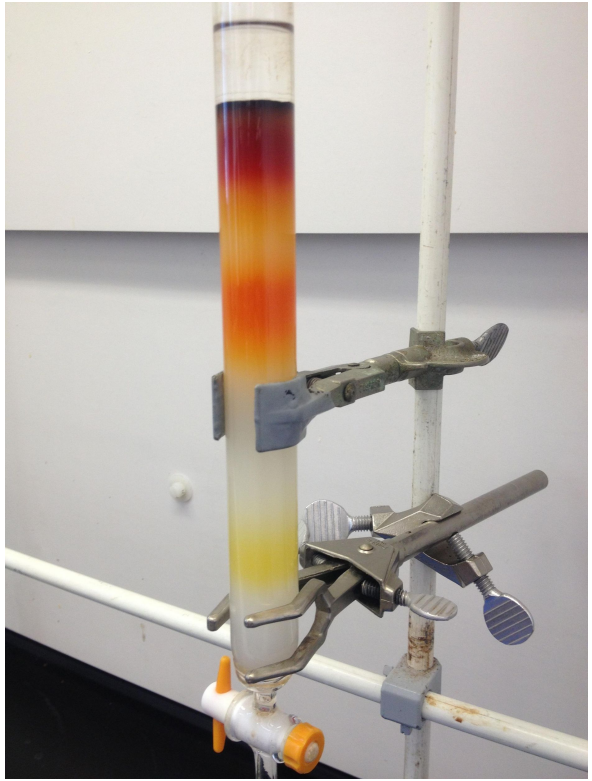
Extractor for **C**hemical **A**nalysis of **Li**pid **B**iomarkers in **R**egolith

1. Building an instrument to extract and purify lipids from regolith to enable optimal GC-MS analysis.
2. Automatically classify lipids as being of ***biotic*** or ***abiotic*** origin.
3. Predict processing parameters for follow-on sampling and sample processing.

Gas Chromatography



Chromatogram



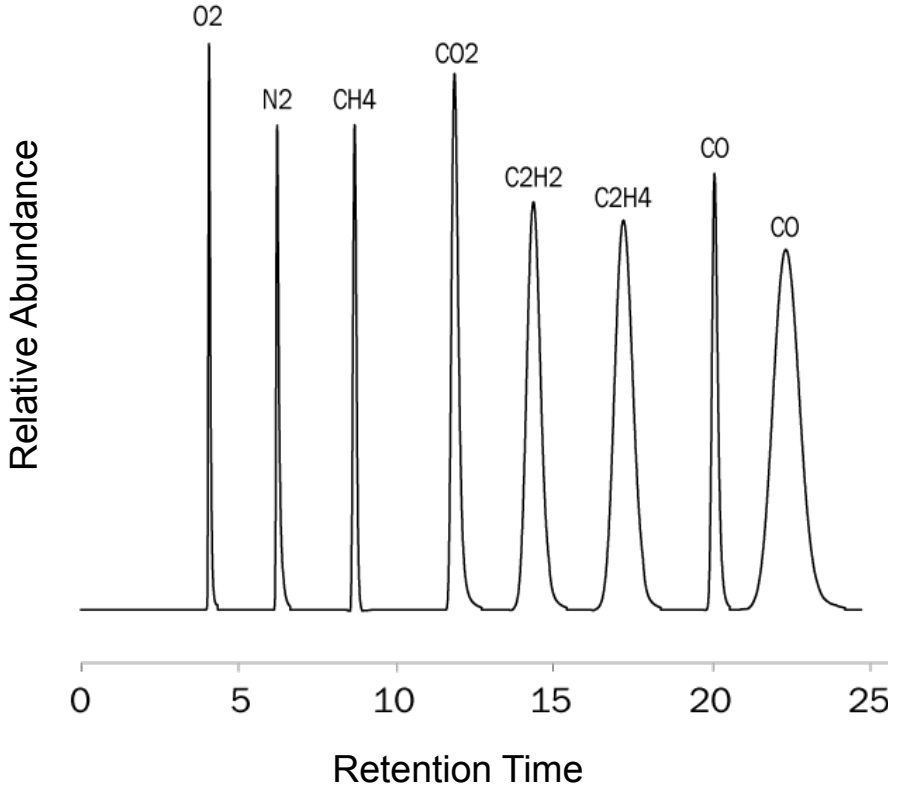


Image source: LDetek

Mass Spectrometry

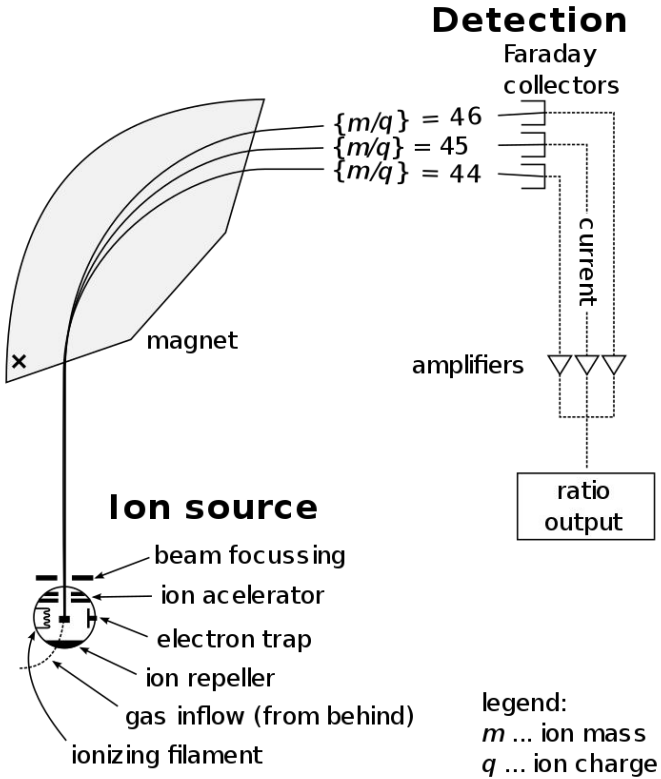


Image source: US Geological Survey via Wikipedia

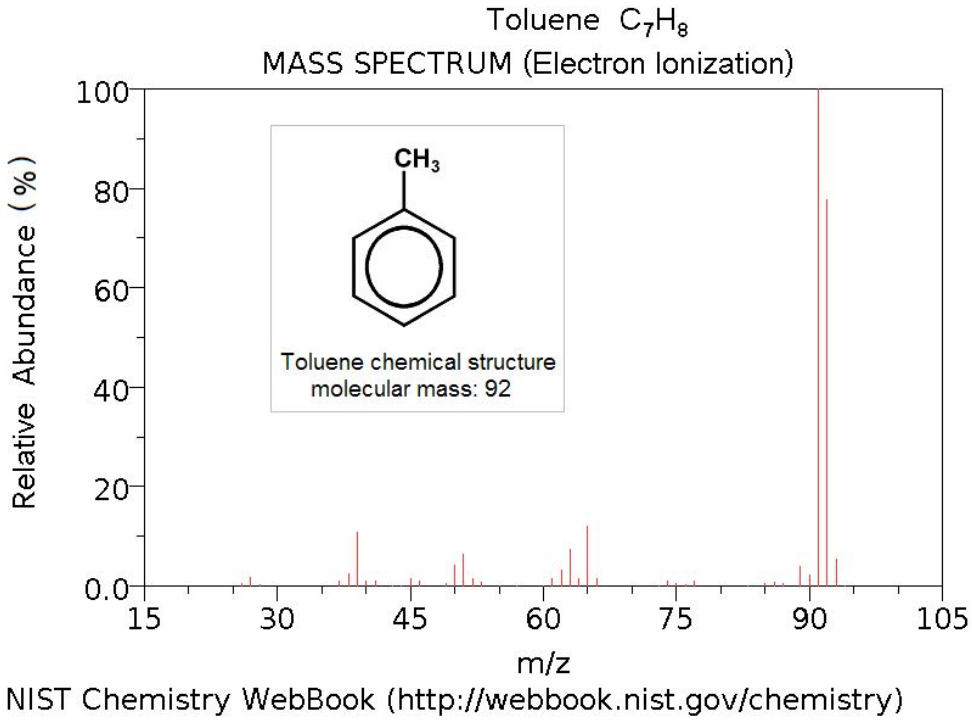


Image source: US Government-NIST via H. Padleckas

Gas Chromatograph - Mass Spectrometer

- Combines both modalities to yield more information about samples.
- Amenable to operations on other planetary bodies.
- Good at detecting organic molecules that tell us about life (e.g. ***lipids***)

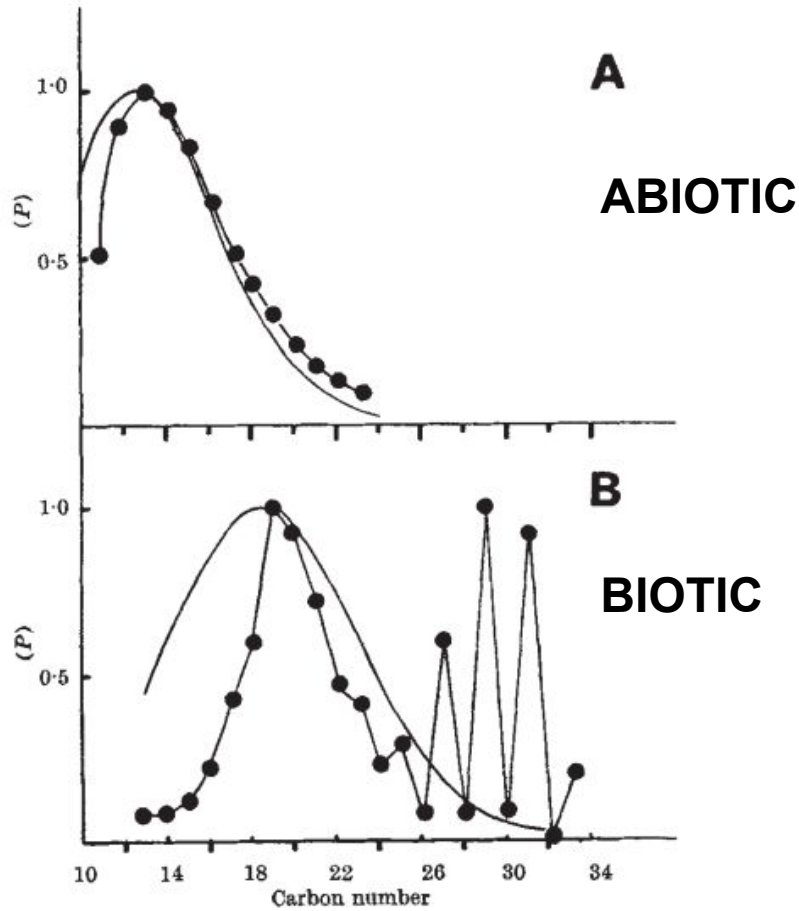
Why Lipids?

Changes the question from:

Can organisms live here?

to:

Does this sample contain an organism?



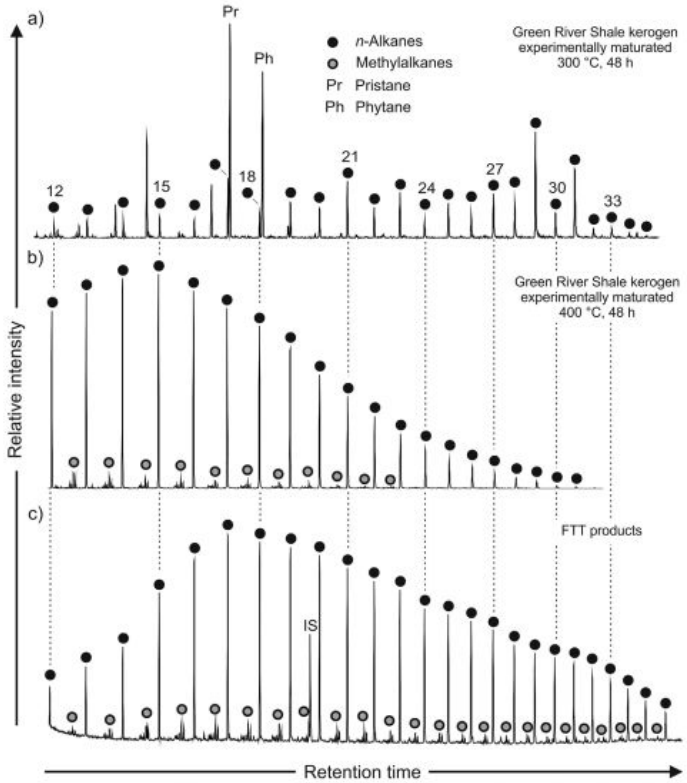
Source: *A Physical Basis for Life Detection Experiments*, Lovelock, Nature #4995 (1965)

Can We Classify Origin Process for Lipids?

BIOTIC

ANCIENT BIOTIC

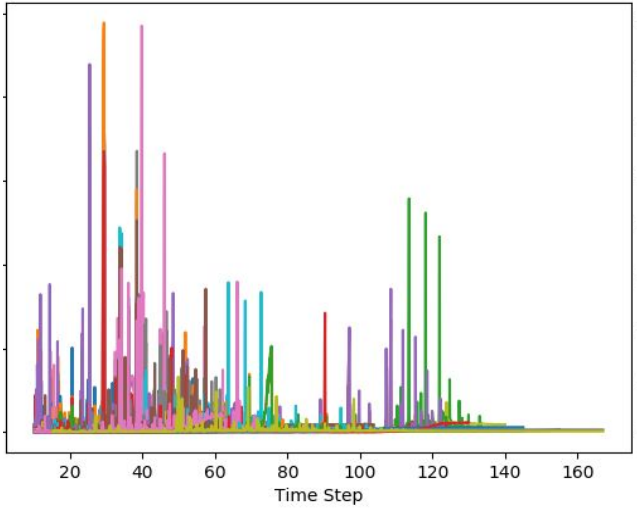
ABIOTIC



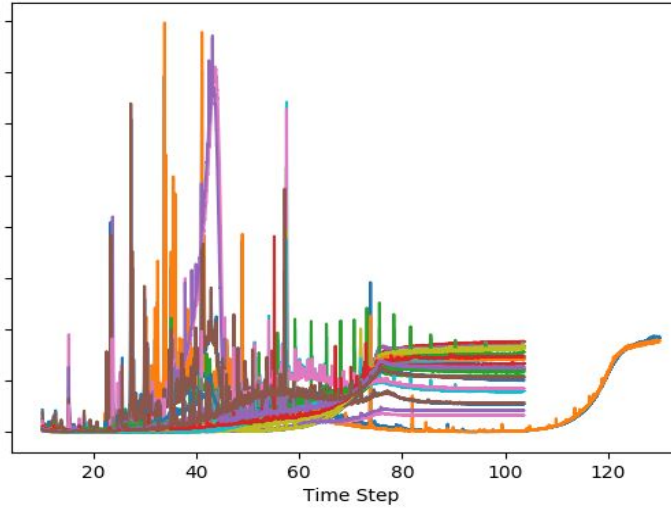
- Not just looking for lipids in old places.
- Geological processes can transform signature of materials of ***biotic*** origin to look like materials of ***abiotic*** origin.

Source: *Assessing the diversity of lipids formed via Fischer-Tropsch-type reactions* Mißbach *et al.*, Organic Geochemistry 119 (2018)

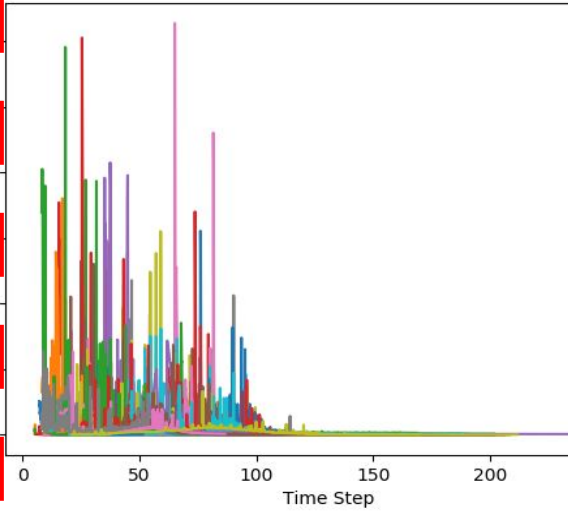
Can We classify Origin Process for Lipids?



BIOTIC

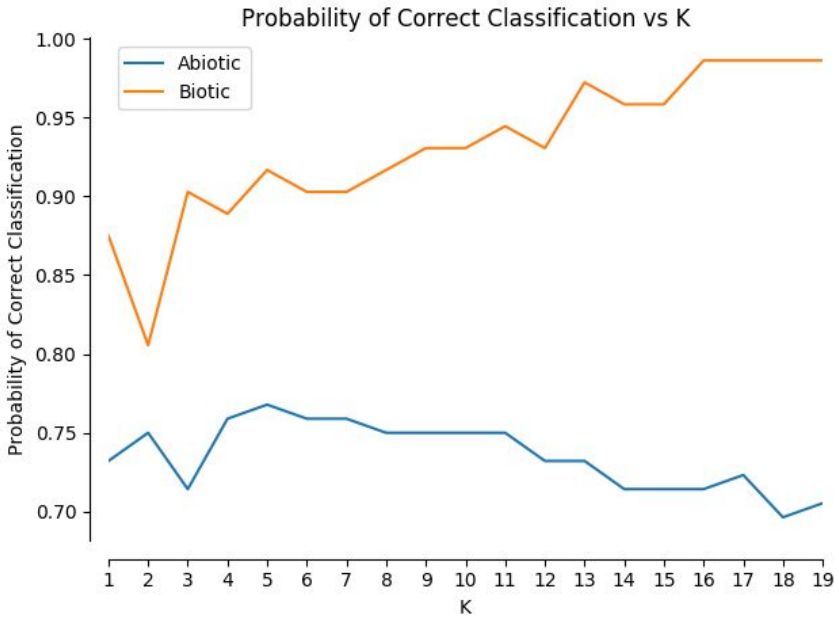


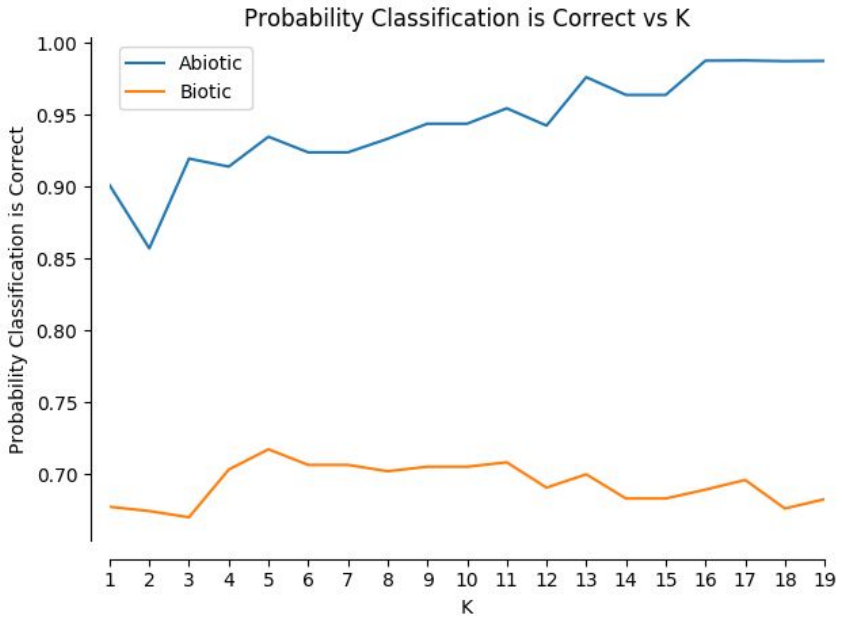
ANCIENT BIOTIC



ABIOTIC

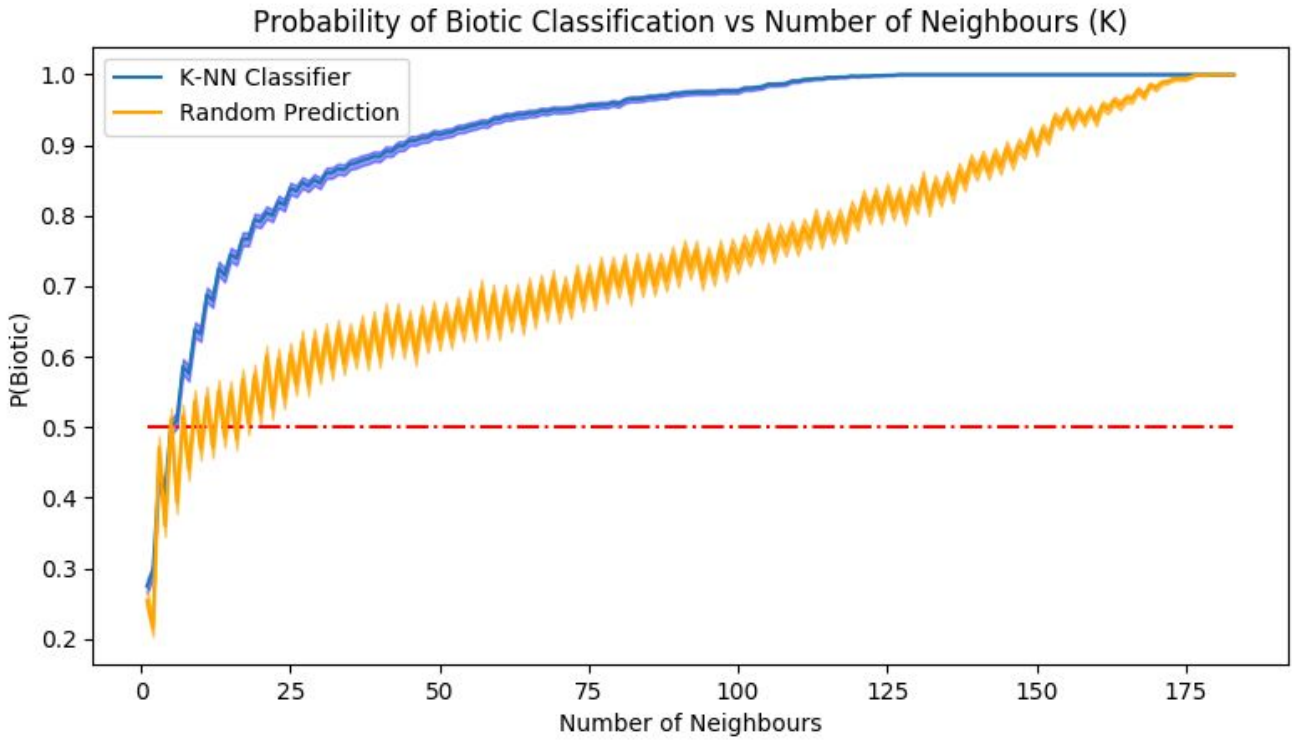
Preliminary Results: Yes





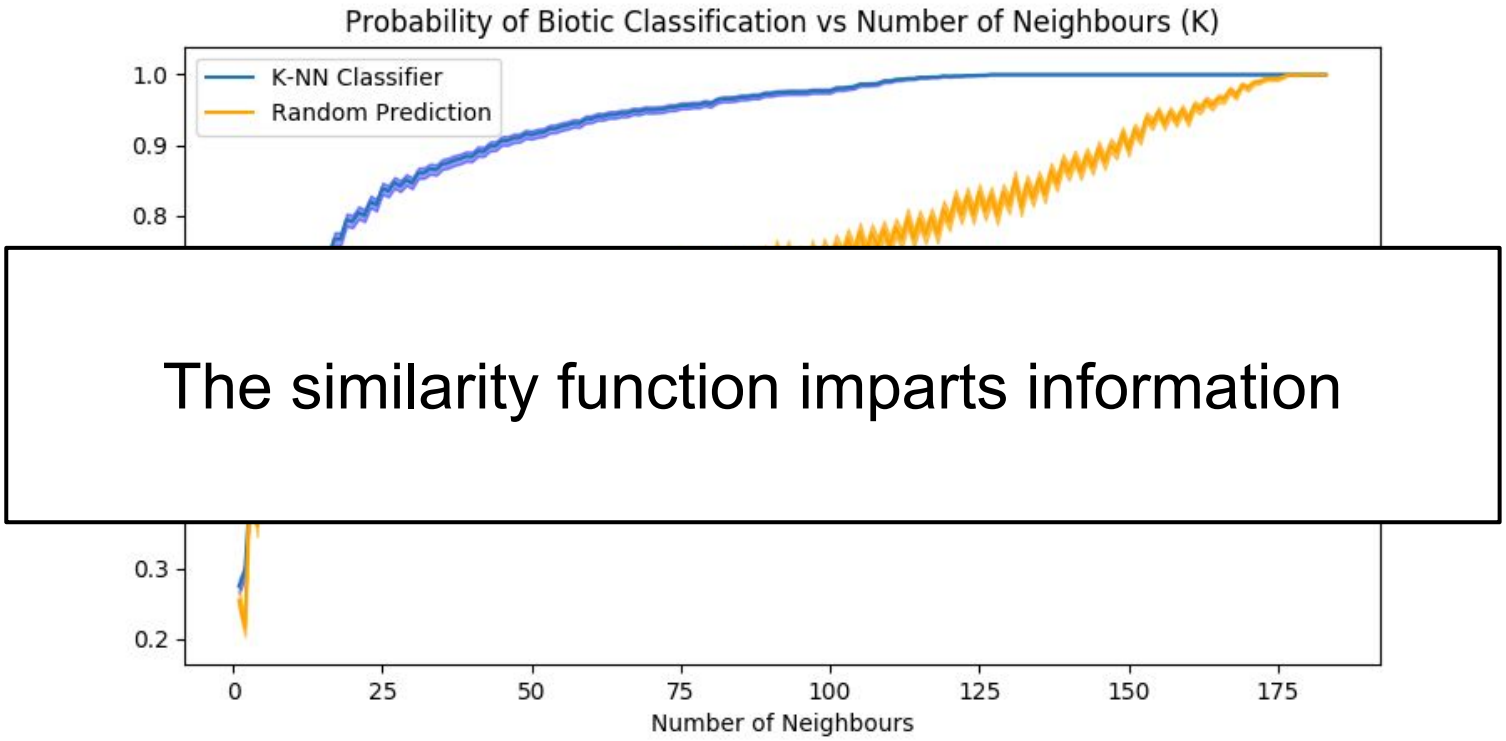
Leave-one-out validation results using K-Nearest Neighbour Classifier

More thrilling results!



Classification of 1931 ***biotic*** samples.

More thrilling results!



The similarity function imparts information

Classification of 1931 ***biotic*** samples.

Continuing Steps

- Improving Dataset quality
  - Existing dataset represents multiple decades of collection.
  - Abiotic samples are exceedingly rare.
- Considering alternative representations of GCMS data

Questions?